5. STANDARD MODEL BACKGROUNDS AND TIMING BIASES IN THE EXCLUSIVE $\gamma + \rlap{/}{E}_T$ FINAL STATE

In this Chapter we describe the SM backgrounds that are present in the exclusive $\gamma_{delayed} + \rlap{/}{E}_T$ final state. We first establish what the backgrounds are to the $\gamma + \rlap{/}{E}_T$ final state and their relative importance to the final search region used in this thesis. As we will see, various SM backgrounds have timing distributions, after the preliminary cuts, that can be very biased toward large values of t_{corr} when the wrong vertex is selected. Moreover, in the exclusive $\gamma_{delayed} + \rlap{/}{E}_T$ final state there is explicitly a lack of other final state particles, thus there are fewer tracks present in the event to produce a vertex. This means that for the SM backgrounds the collision which produced the photon is both less likely to have its vertex reconstructed and less likely to be selected as the highest ΣP_T vertex used in the timing measurement and photon identification.

After a description of the timing distribution for wrong vertex events and detailing the SM backgrounds, we present a study of SM events that give large times and find, in general, three types of events that give large times. These large times are produced two ways, when a background source: 1) produces a small number of individual events with anomalously large times and/or 2) passes the final $\gamma + \rlap{/}{E}_T$ requirements in a way this is biased, on average, towards positive t_{corr} when a wrong vertex is selected. After describing each of these ways that events can have large times, we discuss how to remove/mitigate each problem. Once these steps are taken to reduce the amount of bias present in the wrong vertex distribution we then look at the timing distribution of the known SM backgrounds. In the next chapter we describe how to measure the remaining amount of bias present in the sample.

5.1 Overview of the double Gaussian description of the timing

As described in Chapter 1, the timing distribution for SM events is well described by a combination of events with a right vertex and events where the a wrong vertex is

selected. We next detail more about why our events have these Gaussian shapes. The origin of the Gaussian timing distribution for right vertex events can be understood from the fact that all the measurements of the four quantities for t_{corr} in Equation 1.8 are simple, but dominated by the EMTiming measurement which is a Gaussian measurement (when a discriminator fires a TDC) which has a resolution of 0.5 ns. Taken in quadrature with the other terms (resolution of the vertex time and the time of flight) we get a Gaussian with an RMS of 0.65 ns, as shown in Chapter 3. More details about the resolution of the measurement are given in Reference [52].

The timing distribution for wrong vertex events is more subtle, but still readily described by a Gaussian with an RMS of $\sim$2 ns but a mean that is not zero, as shown in Figure 1.16. To understand this we describe the calculation of the timing distribution for wrong vertex (WV) events. Specifically,

$$t_{corr}^{WV} = t_f - t_0^{WV} - \frac{|\vec{x_f} - \vec{x_0}^{WV}|}{c} = t_f - t_0^{WV} - TOF_{WV} \tag{5.1}$$

where t_0^{WV} and x_0^{WV} are the time and collision z position of the wrong vertex, and where we have relabeled the last term as the time-of-flight from the wrong vertex to the position of the photon in the calorimeter, TOF_{WV}. However, we note that $t_f = t_0^{RV} + TOF_{RV}$ where TOF_{RV} is the time-of-flight from the true collision point to calorimeter location. Plugging into the equation above we find:

$$t_{corr}^{WV} = (t_0^{RV} + TOF_{RV}) - t_0^{WV} - TOF_{WV} \tag{5.2}$$

and re-writing in a suggestive form we find

$$t_{corr}^{WV} = (t_0^{RV} - t_0^{WV}) + (TOF_{RV} - TOF_{WV}) \tag{5.3}$$

The first term in Equation 5.3 is made up of two parts, both given by the Tevatron beam timing structure as described in Section 2.4.5. Both are independent of each

other and well described by a Gaussian with a mean of zero and an RMS of 1.28 ns as shown in Figure 3.10. Thus, we expect them to give a contribution to the final value of $< t_{corr}^{WV} >$ of 0 ns with an RMS of $\sqrt{2} \cdot 1.28 =$ 1.8 ns and a shape that is roughly Gaussian. The second term is pure geometry and can have a mean value that can be many hundreds of ps, but an RMS that, except for the most mis-measured events, is typically small compared to the 1.8 ns of the first term. We will often refer to this term as the bias and it is what produces the events with a large time as well as an overall mean in the wrong vertex distribution. Thus, we see that the wrong vertex timing distribution is expected to not be centered at zero, but will be, to a high degree of approximation, well described by a Gaussian with an RMS slightly above 1.8 ns because of the resolutions of the EMTiming system and the COT vertex time; as we will see in Section 5.6 the RMS will be measured measured with MC and data to be 2.0 ± 0.1 ns taking into account all known effects, such as the 0.5 ns resolution on the EMTiming. As was shown in Figure 3.2, a $< t_{corr}^{WV} > \sim 0.5$ ns can cause there to be twice as many events in the signal region as in the control region for SM backgrounds.

This description of the t_{corr} distribution as a double Gaussian with right and wrong vertex is confirmed when we look at data from $e + \rlap{/}{E}_T$ calibration events in the top of Figure 5.1 (data selection described in Table 3.1). Here we pick the highest ΣP_T vertex as the inital collision position (t_0 and z) to calculate the corrected time. Using the matching of the electron track to the vertex we can further divide this distribution into the right and wrong vertex samples individually. This is done in order to verify the Gaussian nature of both distributions. This description clearly models the data very well and thus gives us confidence in the use of a double Gaussian function to describe the corrected time of collision events.
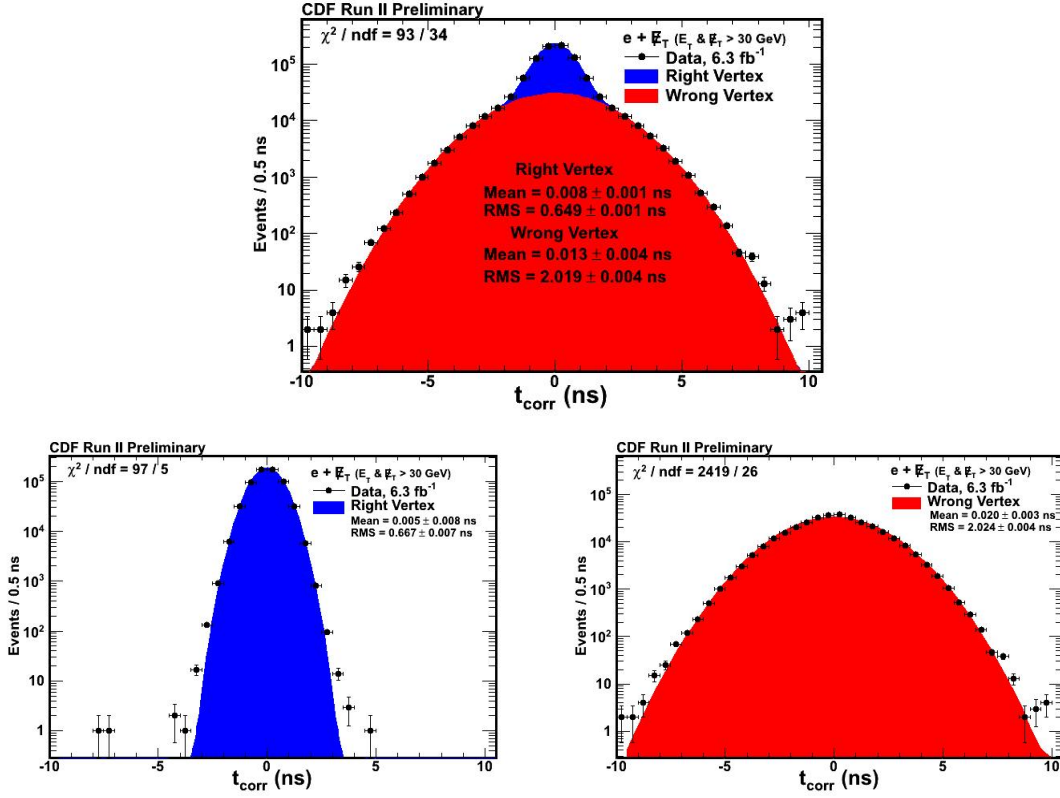
Fig. 5.1. The t_{corr} distribution using the $e+\not{E}_T$ calibration data (defined in Table 3.1) using the highest ΣP_T vertex (top) as well as the right and wrong vertex Gaussians using the matching of the electron track to the vertex (bottom) verifying the description of the timing distribution as being well described by a double Gaussian distribution of two well understood and seperate contributions.

5.2 Standard Model Backgrounds

We now turn to the various SM processes that can produce the exclusive $\gamma+\not{E}_T$ final state with an eye towards the biases each might bring to the wrong vertex timing distributions. Since the appearance of a single photon plus missing energy with little other activity present in the detector is a very unlikely thing to have coming directly from SM processes, the presence of these backgrounds normally results from a coincidence of various processes taking place in the detector which is

likely to have a bias. For example, a W$\rightarrow e\nu$ event will usually produce significant $\not{E}_T$ from the neutrino, but will only have the electron mis-reconstructed as a photon a tiny fraction of the time, but in a way that produces a biased time. This is because electrons that have longer path lengths are more likely to produce a fake photon as well as have a large value of t_{corr}^{WV} since TOF_{RV} will be longer, on average, than TOV_{WV}. Broadly speaking this type of coincidence of unusual occurances is what occurs for all our SM processes. Almost all have a very large production cross-section while the occurrence of the detector or reconstruction failing to properly reconstruct the collisions is a small fraction. This multiplication of a large number (SM cross-section) times a small number (detector/reconstruction failure) makes a direct estimate of SM backgrounds using only MC methods difficult as elaborated upon below.

We focus on MC simulation of all known SM sources that produce $\gamma+\not{E}_T$ for any and all types of example events that could give a large time bias. Said differently, we use each observed event in the MC with a large time to try and construct an understanding of what types of events cause a bias. By studying all the sources below, we have confidence that we have considered all the important sources of large time events that can be produced.

We next describe all the SM backgrounds that produce the $\gamma+\not{E}_T$ final state, summarized in Table 5.1 using the MC samples described in Table 2.12. We note that the rates from each background is never determined individually, but are done collectively using the data-driven procedures described in Chapter 6.

- W$\rightarrow$ e$\nu \rightarrow \gamma_{fake}+\not{E}_T$:

  The first of the processes considered for the exclusive $\gamma_{delayed}+\not{E}_T$ final state comes from the SM process where a W boson is produced from the collision and the subsequently decays to an electron (e) and a neutrino (ν) but where the electron is identified as a photon (commonly referred to as a γ_{fake}) in the detector. The probability of the electron faking a photon of $\sim$1% [80]. Of

Standard Model Process
W$\rightarrow$ e$\nu \rightarrow \gamma_{fake} + \not{E}_T$
γ+Jet $\rightarrow \gamma$+Jet$_{Lost} \rightarrow \gamma + \not{E}_T^{fake}$
Z$\gamma \rightarrow \gamma\nu\nu \rightarrow \gamma + \not{E}_T$
W$\rightarrow \mu\nu \rightarrow \gamma_{fake} + \not{E}_T$
W$\rightarrow \tau\nu \rightarrow \gamma_{fake} + \not{E}_T$
W$\gamma \rightarrow$ lepton$_{lost}\ \nu\gamma \rightarrow \gamma + \not{E}_T$

Table 5.1

Summary of the various SM backgrounds considered for the exclusive $\gamma_{delayed}+\not{E}_T$ final state. Each sample of events is simulated with MC with more details given in Table 2.12.

particular importance, W$\rightarrow$ e$\nu \rightarrow \gamma_{fake}+\not{E}_T$ events often have the wrong vertex selected, for reasons explained later, and do so in ways that have large values of $TOF_{RV}-TOF_{WV}$ (the second term in Equation 5.3). Specifically, the ways that electrons fake photons, and the E_T distribution of electrons from W$\rightarrow$ e$\nu \rightarrow \gamma_{fake}+\not{E}_T$ events can both bias $< t_{corr}^{WV} >$ distribution as well as significantly change which events enter and leave our sample. This is elaborated on further in Section 5.5.1. We note that W$\rightarrow e\nu$ events that are produced in association with a photon, but where the electron is not reconstructed by the detector (referred to as Wγ) are included in a later background.

- γ+Jet $\rightarrow \gamma$+Jet$_{Lost} \rightarrow \gamma + \not{E}_T^{fake}$:

QCD processes, such as a photon plus a jet coming from a quark or gluon (γ+Jet), do not have any intrinsic $\not{E}_T$ in their production. However, a large value of $\not{E}_T$ can be measured in the event via mis-reconstruction or mis-measurement of the energy contained in the reconstructed jet or photon. One such example of how this can occur is when energy from the jet is produced at large $|\eta|$ and travels down the beam pipe or otherwise hits an un-instrumented region of the detector. While the fraction of QCD events with a mis-measurement

of $\not{E}_T$ as high as 45 GeV is very small ($<5\%$), the total QCD cross section is very large ($\mathcal{O}(1000$'s pb$)$) [3], thus making this another potentially large background to $\gamma+\not{E}_T$ final state. In particular, it can be a significant background to exclusive $\gamma_{delayed} + \not{E}_T$ because the topology of the events that do pass have a high correlation between large $\not{E}_T$ and the vertex being mis-measured or not reconstructed in a way that biases this sample towards large values of $TOF_{RV} - TOF_{WV}$. This process is presented in greater detail in Section 5.5.3.

- $Z\gamma \rightarrow \gamma\nu\nu \rightarrow \gamma + \not{E}_T$:

An irreducible background in the exclusive $\gamma+\not{E}_T$ final state comes from the production of a Z boson made in association with a photon from initial state radiation. The Z subsequently decays into to a pair of neutrinos which escape the detector in the form of $\not{E}_T$ and we thus end up with the $\gamma+\not{E}_T$ final state. While the decay rate of the Z boson to pairs of neutrinos is much higher than to charged leptons (which we are able to veto with a high efficiency), the rate at which a 45 GeV photon is produced in the central part of the detector is small. The topology of these events causes a relatively small wrong vertex mean, as will be discussed further in Section 5.6. This further diminishes the importance of this otherwise irreducible background.

- $W \rightarrow \mu\nu, W \rightarrow \tau\nu$, and $W\gamma$ sources:

The last major sources of backgrounds in the exclusive $\gamma+\not{E}_T$ final state are from W$\rightarrow$ lepton $+ \nu$ boson sources where the lepton may fake a photon or become lost. Two such examples of this decay are W$\rightarrow \mu\nu \rightarrow \gamma_{fake} + \not{E}_T$ and W$\rightarrow \tau\nu \rightarrow \gamma_{fake} + \not{E}_T$. While the production cross-section of these processes is known to be relatively large ($\sim$1000 pb [81]), the likelihood of these leptons faking a photon is much smaller than for electrons [61] in the CDF detector. Therefore we do not devote any specific cuts to rejecting these processes. The last of the W boson processes we consider comes from the production of a

W boson with initial or final state radiation where the W boson decays to a lepton plus a neutrino and the lepton becomes "lost" (in contrast to "faking") in the detector giving us the $\gamma + \not{E}_T$ final state. This decay is dealt with in part in Section 5.5.2. In addition we note that the tracking efficiency in the CDF detector is known to be $> 95\%$ [50], thus making this a low rate background in this analysis.

5.3 Presamples and Control Samples for use in the Exclusive $\gamma_{delayed} + \not{E}_T$ Final State

Having finished an overview of the individual SM backgrounds that can produce the exclusive $\gamma + \not{E}_T$ final state, we now turn our attention to defining a number of "presamples" and "control samples" to help study them. A presample is a set of events that pass a set of selection requirements, but not all the final requirements so that we may study the effect of some of the later requirements. A control sample is a sample of events, data or MC, that is selected in a way so that it can be compared to the results expected from $\gamma + \not{E}_T$ events in data after all the requirements.

As previously mentioned in Table 5.1, we have six sources of SM backgrounds which are studied separately using independent MC data sets whose production were described in Table 2.12. Additionally, we have an e+$\not{E}_T$ data samples that will be used to generate sets of presamples and control samples. The specific set of selection requirements each presample will be given later, for now we simply summarize the two different types of presamples we will use.

- **Exclusive Electron and Missing Energy Sample:**

  This presample will mirror the $\gamma + \not{E}_T$ final state used in the final search except that instead of a photon we require an electron. Since the electron tracks in these events are excluded from use in the SpaceTime vertexing algorithm this presample provides a very good analog to the $\gamma + \not{E}_T$ sample as well as a data-

based way of testing the background estimation methods described in Section 6.1. Particularly, the use of the electron track seperate from the vertex will allow us to have important information about the initial position and time of the event as well as provide a testing ground for our analysis. This presample will be used to create control samples in both our MC and data. The exclusive e+$\not{E}_T$ presample selection requirements is summarized in Table 5.2.

Exclusive $e+\not{E}_T$ Presample Event Selection
Pass Trigger and Good Run List
(See Table 2.2 and 2.3 and Section 2.4)
Pass Electron requirements w/ $E_T^0 > 45$ (30) GeV and $\not{E}_T^0 > 45$ (30) GeV
(See Table 2.9 and Section 2.4.6)
Pass Beam Halo Rejection
(See Table 4.3)
Pass Track Veto for Tracks with $P_T > 10$ GeV
(See Table 2.7)
Pass Jet Veto for Jets with $E_T^0 > 15$ GeV
(See Table 2.5)

Table 5.2

List of selection requirments summarizing the exclusive $e+ \not{E}_T$ presample. Note, that our sample requires one with E_T^0 and $\not{E}_T^0$ at 30 GeV, but we will frequently make a subsample of this with requirements at E_T^0 and $\not{E}_T^0$ at 45 GeV. Note that the trigger, beam halo, and good run list requirements are only for data and not used on the MC.

- **Exclusive Photon and Missing Energy Sample:**

  This presample is designed to mirror the selection used in the preliminary result from 2008 and allow us to study the effects of various background processes in this final state. This set of presample selection requirements will be used to create control samples from our MC datasets, help us understand the sources of biased events, and determine our background rejection requirements. The exclusive $\gamma+\not{E}_T$ presample selection requirements are summarized in Table 5.3.

Exclulsive $\gamma + \not{E}_T$ Presample Event Selection
Pass Trigger and Good Run List
(See Table 2.2 and 2.3 and Section 2.4)
Pass Tight Photon requirements w/ $E_T^0 > 45$ GeV and $\not{E}_T^0 > 45$ GeV
(See Table 2.8 and Section 2.4.6)
Pass Beam Halo Rejection
(See Table 4.3)
Pass Cosmics Rejection
(See Table 4.2)
Pass Track Veto for Tracks with $P_T > 10$ GeV
(See Table 2.7)
Pass Jet Veto for Jets with $E_T^0 > 15$ GeV
(See Table 2.5)

Table 5.3

List of selection requirments summarizing the exclusive $\gamma + \not{E}_T$ pre-sample. Note that the trigger, good run list, beam halo, and cosmics requirements are only for data and not used on the MC.

We create ten control samples that will be used throughout the remainder of this thesis. In particular, we create the six $\gamma + \not{E}_T$ samples from the MC datasets described in Table 5.1 and require each event to pass all the preseample requirements in Table 5.3. Similarly, we create four samples of e+$\not{E}_T$ using real data and MC using the requirements in Table 5.2, but where we have one sample with $E_T > 45$ GeV to fully mimic the final requirements, and one with $E_T > 30$ GeV to give us better statistics. Finally, we will create a presample of $\gamma + \not{E}_T$ events from data that will be used to select cosmics events in Section 5.5.3. We will detail the control samples used in Chapter 6 after we have finished the event rejection requirements in this chapter.

5.4 Sources and Categorizing the Causes of Large Mean Shifts

With these samples in hand we are now ready to return to the task of showing the various pathologies and biases present in the exclusive $\gamma + \not{E}_T$ final state. This

will show just how bad the underlying assumption that the wrong vertex mean is centered at $t_{corr} = 0.0$ ns can be and why it was particularly poor for the case in the preliminary study performed in 2008. We next provide a convenient categorization of the dominant effects that cause SM events to have large values of t_{corr}^{WV} as a prelude to a set of selection criteria to reject the most biased effects and minimize the $< t_{corr}^{WV} >$ distribution.

Rather than describe a historical summary of how we came to understand these effects, we quickly summarize the effects that cause large values of large time events and outline criteria to systematically remove or minimize their impact. In particular, the types of events that produce large values of $TOF_{RV} - TOF_{WV}$.

- **Events with Geometric and Kinematic Biases:**

  Incorrect selection of the vertex causes an incorrect $\sin\theta$ to be assigned to the measurement of the photon's E_T. This can cause events to be incorrectly included in our sample, as described in Section 5.5.1. This same effect can also cause a mis-measured/biased timing distribution because, as we will see, the E_T as mis-measured from the wrong vertex is highly correlated with $TOF_{RV} - TOF_{WV}$. This timing mismeasurement can lead to events migrating into the signal region and out of the control region and further exasperate problems in estimating the mean of the wrong vertex, thus 'faking' a signal.

  In order to reduce the impact of this migration of events we will redefine our E_T variable. In doing so we will take advantage of the fact that on average most collisions occur at $z = 0$. This will be described in more detail in Section 5.5.1. This definition will have further advantages when we describe the background estimation techniques in Chapter 6.

- **Events with $e \rightarrow \gamma_{fake}$ Sources:**

  Incorrect selection of the vertex also causes the standard photon identification variables used to reject sources which "fake" a photon in the detector to become

less powerful [82] because they assume the primary vertex. Specifically, identifications like track isolation currently require that the tracks being considered are within 5 cm of the primary vertex. Since selection of the correct vertex becomes a problem in the exclusive $\gamma + \not{E}_T$ final state, it becomes necessary to employ new 'fake' photon rejection. A second, and correlated, issue is that the same thing that makes an electron fake a photon, e.g. large path length, is also correlated with a large value of $TOF_{RV} - TOF_{WV}$ causing electrons that fake photons have a large value of $< t_{corr}^{WV} >$.

Since the typical tools for $e \rightarrow \gamma_{fake}$ are found to be not as powerful in the case when the wrong vertex is selected, they are a problem for the exclusive $\gamma_{delayed} + \not{E}_T$ analysis. For this reason we have developed a new method whereby we can reject 67% of events coming from W$\rightarrow e\nu \rightarrow \gamma_{fake} + \not{E}_T$ with a 95% efficiency for real photons. This removes the majority of the most biased events from the $\gamma + \not{E}_T$ sample. This is detailed in Section 5.5.2.

- **Events originating from Large $|z|$:**

  Finally, the last of the important reconstruction pathologies that effects the timing distribution comes from events originating from a large $|z|$ collision location. If there is a collision that occurs with $|z| >$60 cm which creates a real photon that is then observed in the central calorimeter, this vertex will not be reconstructed or reported by the SpaceTime vertexing algorithm. This is due to the fact that we explicitly require that the SpaceTime vertices used in the exclusive $\gamma + \not{E}_T$ analysis to come from $|z| <$60 cm in order to have a good timing measurement associated with the vertex. If there happens to be a min-bias event near the center of the detector that creates a good SpaceTime vertex, a wrong vertex will be assigned to the event. This is most easily seen in $\gamma +$ jet events where there is a high degree of correlation between the production of fake $\not{E}_T$ and large values of $|z|$.

Since the timing bias from wrong vertex events can be especially large if the true collision occurred with $|z| > 60$ cm, we reject events that have evidence of a collision occurring at large $|z|$ position. This will help minimize the wrong vertex mean bias as well as reduce the most pathologically mis-reconstructed events that have the largest value of $TOF_{RV} - TOF_{WV}$. This veto is detailed in Section 5.5.3.

5.5   Rejecting Backgrounds with Large Times

With this basic understanding of the types of events that have large times from wrong vertex pathologies, we next move to a description of how to reject many of these events. Since they come in three basic types, have created a set of three rejection criteria to systematically remove or minimize their impact. We will discuss in more detail how to measure the remaining amount of bias in the next chapter.

5.5.1   Minimizing the Correlations Between Geometric and Kinematic Biases

As seen in Figure 5.2, selection of a wrong vertex affects both the t_{corr} calculation as well as the E_T calculation. This occurs because the geometry of the events effects the kinematics, and vice versa and, for some samples of events, can cause events selected using the E_T of the photons to bias the mean time of the t_{corr}^{wv} distribution. To understand why this is particularly important for this search, we next describe the correlation between the measured E_T and t_{corr}.

Before we begin this discussion it is important to know the standard ways of measuring E_T in the detector which will illustrate why we do something a little different, as first described in Chapter 2. Specifically when we select the highest ΣP_T vertex we calculate $E_T^{Measured}$ and $t_{corr}^{Measured}$, but these may not be the same values as when measure them from the true vertex, E_T^{True} and t_{corr}^{True}. The same geometric effects which cause $E_T^{Measured} > E_T^{True}$ can also cause $t_{corr}^{Measured} > t_{corr}^{True}$. Before going

forward it is also worth clarifying the use of the angle measured from the beamline, θ. As we can see from Figure 5.2, there are many different θ's that are possible to define. For example, there is θ_{True} which is θ measured from the true collision, and can be referred to as θ_{RV}. If we select a wrong vertex then we could refer to the $\theta_{Measured}$ as θ_{WV}. In general $\theta_{Measured}$ can be either θ_{WV} or θ_{RV}. Previously in this thesis we have defined θ_0 as measured from the center of the detector and θ_{vertex} as measured from the vertex. In the following discussion we will use $\theta_{Measured}$ and θ_{True} as our notation to be more explicit where it is most helpful.

$$E_T^{True} = E\sin(\theta_{True})$$
$$E_T^{Measured} = E\sin(\theta_{Measured})$$

CDF Calorimeter
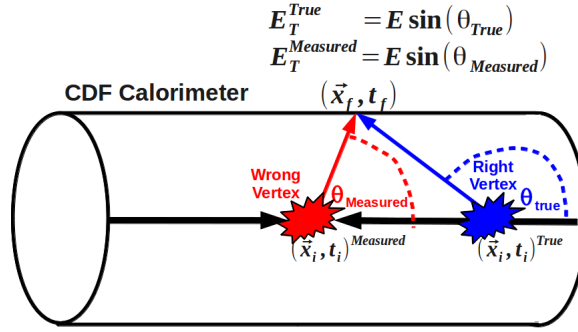
Fig. 5.2. A schematic drawing of a W$\rightarrow e\nu \rightarrow \gamma_{fake} + \not{E}_T$ where we have selected a wrong vertex. For this example, the selection of the wrong vertex leads to an $\theta_{Measured} > \theta_{True}$ where θ_{True} is the real angle the photon/electron came from. This results in a larger measured value for E_T ($E_T^{Measured} > E_T^{True}$), thus preferentially causing us to select these events. Furthermore, the path length calculated for the wrong vertex is shorter than the true path length of the photon/electron resulting in an apparent longer time of flight and thus $t_{corr}^{Measured} > t_{corr}^{True}$ causing an overall shift in the mean.

Consider the configuration in Figure 5.2 where the vertexing algorithm does not select the correct collision point, either because it is not reconstructed or because a higher ΣP_T vertex from a min-bias interaction happens to exist. In this case we have $\theta_{Measured} > \theta_{True}$, so that $TOF_{WV} < TOF_{RV}$ ($|\vec{x}_f - \vec{x}_{Measured}| < |\vec{x}_f - \vec{x}_{True}|$),

resulting in $t_{corr}^{Measured} > t_{corr}^{True}$ (ignoring the contribution from t_i and t_{WV}), or said differnetly $(TOF_{RV} - TOF_{WV}) > 0$ ns.

At the same time, since $\theta_{Measured} > \theta_{True}$, we find $E_T^{Measured} > E_T^{True}$. This implies that events that have a positively shifted $t_{corr}^{Measured}$ will also have a larger $E_T^{Measured}$. This fact has a remarkable consequence. Namely, this means that some of the events that have E_T^{True} slightly less than 45 GeV, and should not be in our sample of events, can have an $E_T^{Measured} > 45$ GeV and will enter the sample because we chose the wrong vertex. Since these events will also have a timing bias, this means that all the events that enter the sample (i.e., that pass the cuts but shouldn't have) will have $< t_{corr}^{WV} > > 0$ ns.

The converse is also true, a configuration with a mis-measured vertex, where $\theta_{Measured} < \theta_{True}$, would lead to a lower measured of $t_{corr}^{Measured}$ and lower value of $E_T^{Measured}$. Specifcally, events that have E_T^{True} slightly more than 45 GeV and should remain in our sample of events but have an $E_T^{Measured} < 45$ GeV will leave the sample because of the choice of the wrong vertex. These events have a negitive t_{corr} timing bias, but go unobserved since they leave the sample. The bottom line of all this is that misidentification of vertices leads to values of $t_{corr}^{Measured}$ and $E_T^{Measured}$ being shifted in the same positively biased direction.

Tying this all together we find that events that migrate into the sample have large times and events that leave the sample have smaller times. While this might not be a big effect in principle, the number of events entering and leaving around an E_T cut is frequently asymmetric as demonstrated in Figure 5.3 which shows our presample of W$\to e\nu$ MC events that pass the requirements in Table 5.3, but have lowered the E_T requirement, E_T^{cut} to 25 GeV so we can see the full E_T distribution. Since we are placing our E_T^{cut} just past the peak of E_T^{True}, many events make it past the value of E_T^{cut}, resulting in a higher value of $< t_{corr}^{WV} >$ and many events just fail the cut also resulting in a higher value of $< t_{corr}^{WV} >$. At 45 GeV (where our E_T

selection resides) the peak of the E_T^{True} distribution makes the migration effect very significant.
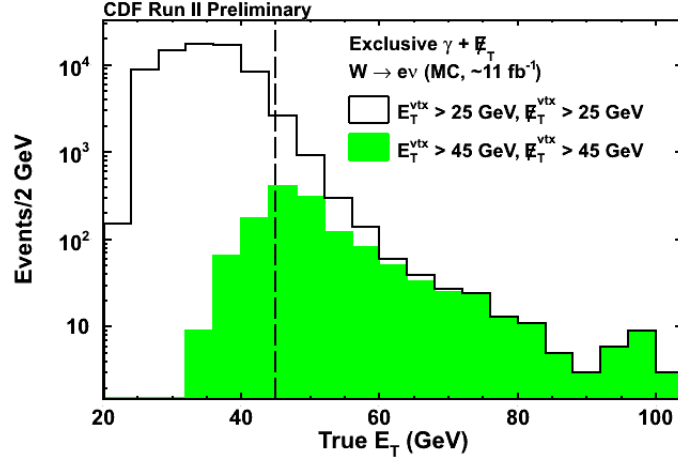

Fig. 5.3. The E_T^{True} distribution for a sample of W$\rightarrow e\nu$ MC events selected with Table 5.3 that make it into the exclusive $\gamma + \not{E}_T$ presample but with two different E_T requirements. The unshaded histogram is the true E_T for electrons that fake photons with $E_T^{True} >$25 GeV while the solid histogram (shown in green) is the E_T^{True} for electrons that were identified as photons, and passed the $E_T^{Measured} >$45 GeV cut. Events both entering the sample (green events below the dashed line) and the events leaving the sample (area under the open histogram but above the green, to the right of the dashed line) bias the measurement of t_{corr}^{WV} since large time events enter the sample and low time events leave the sample.

In order to mitigate this effect, we exploit the fact that at CDF most collisions occur on average at $z = 0$. Thus, if we instead define E_T for the photons reconstructed in our events from $z = 0$, E_T^0, (as was used in Tables 5.2 and 5.3) instead of from the highest ΣP_T vertex, we will never be exactly right on an event-by-event basis, but be more right on average for events when we select the wrong vertex. This is particularly important as it reduces the amount of bias for the most biased events. The effect of this is that fewer events will be "promoted" into our sample by having

$E_T^{Measured} > E_T^{True}$ (on average) as well as fewer events being "demoted" out of our sample with $E_T^{Measured} < E_T^{True}$. Similarly this means that fewer events with $t_{corr}^{Measured} > t_{corr}^{True}$ will be entering our sample on average as well as fewer events with $t_{corr}^{Measured} < t_{corr}^{True}$ leaving our sample. Perhaps more importantly, the variation of the most extreme events in $TOF_{RV} - TOF_{WV}$ is smaller than $TOF_{WV} - TOF_0$. The net result is the kinematic bias that was present before as a result of this definition of E_T is diminished and the most biased events are removed.

To see how much this redefinition can help, we consider our presamples of $e + \not{E}_T$ from both data and W$\rightarrow e\nu \rightarrow e + \not{E}_T$ MC selected using the cuts in Table 5.2. The results are shown in Figure 5.4 where we select events based on $E_T^{Measured}$ and $E_T^0 > 45$ GeV. On the left hand side of Figure 5.4 we see that in both data and MC have $< t_{corr}^{WV} > \sim 0.4$ ns when we select on $E_T^{Measured}$. On the right hand side of Figure 5.4 we find $< t_{corr}^{WV} >$ is only ~ 0.2 ns when select on E_T^0 and $\not{E}_T^0$. The remarkable agreement between data and MC gives us great confidence that the understanding of the source of this bias is well modeled and understood, and that a large portion of the bias present in the wrong vertex distribution for SM processes in the exclusive $\gamma + \not{E}_T$ final state can be reduced. It is also worth noting at this point that the timing distributions for both data and MC are well described by a double Gaussian distribution, as expected.

5.5.2   Rejecting Events from $e \rightarrow \gamma_{fake}$ Sources

Since the W$\rightarrow e\nu \rightarrow \gamma_{fake} + \not{E}_T$ process is a considerable background to any search with a $\gamma + \not{E}_T$ final state it is useful to understand why this background gives a large timing bias and what causes this background to occur so we can reject it more effectively. This discussion follows the description given in Reference [83]. The first point is that, in the exclusive $\gamma + \not{E}_T$ final state, the lack of other activity in the detector required means the primary vertex is both less likely to be reconstructed and less likely to be the highest ΣP_T vertex in the event. Additionally, as discussed
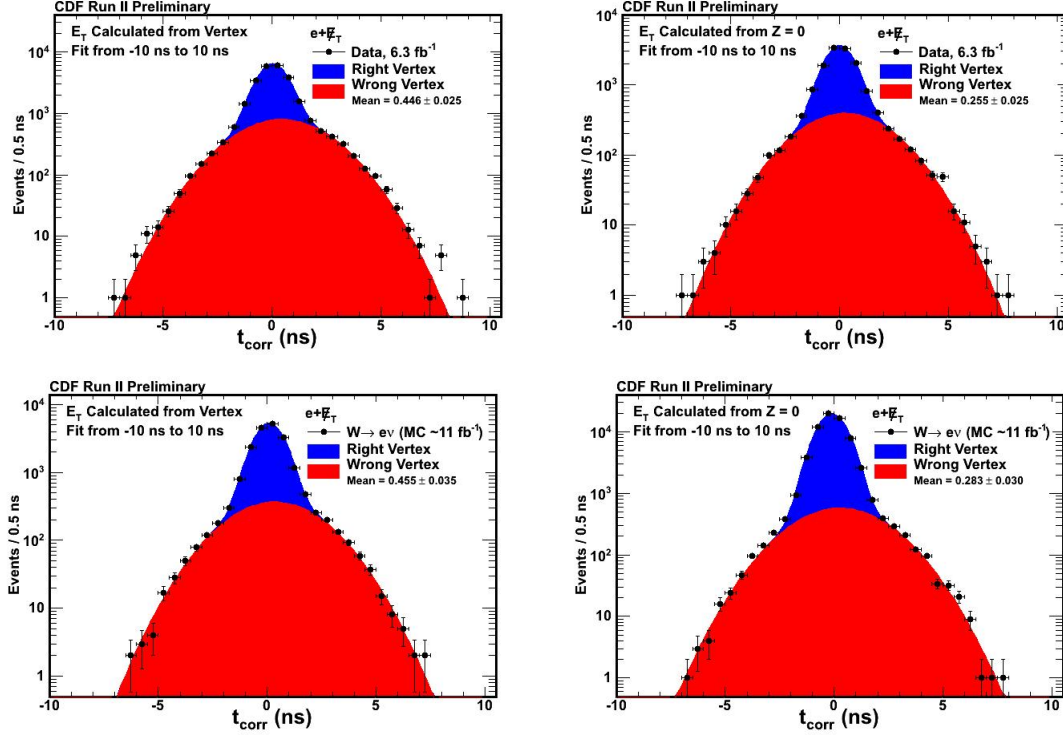
Fig. 5.4. The t_{corr} timing distribution for the e+$\not{E}_T$ presamples described in Table 5.2 from data (Top Row) and W$\rightarrow e\nu$ MC (Bottom Row) when the events are selected using $E_T^{Measured} >$45 GeV (LHS) and using $E_T^0 >$45 GeV (RHS). This showes that you can reduce the value of $< t_{corr}^{WV} >$ by simply calculating E_T and $\not{E}_T$ from $z = 0$. Note that data and MC give very similar results showing that this effect is well understood.

in Section 5.4, the same geometric effects that lead to a positive time bias are also the same issues that contribute electrons to fake photons. Namely an electron with a longer path length from the collision point to the calorimeter is more likely to 'fake' a photon. This is because as the electron traverses more material it is more likely to interact and bremsstrahlung a large energy photon that will be identified in the calorimeter as good photon. A longer path length for wrong vertex events implies $TOF_{WV} < TOF_{RV}$ ($|\vec{x}_f - \vec{x}_{Measured}| < |\vec{x}_f - \vec{x}_{True}|$), which also implies

$E_T^{Measured} > E_T^{True}$ and $t_{corr}^{Measured} > t_{corr}^{True}$ and thus also is subject to the same bias arguments just given.

To reduce the fraction of the time an electron fakes a photon we first note that the dominant way that electrons fake photons is when there is a hard bremsstrahlung interaction in the detector material. A schematic drawing of this process of how these $e \to \gamma_{fake}$ events from hard bremsstrahlung interactions appear in the detector is shown in Figure 5.5. As an electron travels through the detector material, the hard interaction can cause it to lose a large fraction of its energy to a photon. The electron's trajectory is severely affected by the energy and momentum loss; it may either leave a much lower energy deposition in a calorimeter or be swept away completely by the magnetic field of the solenoid. The bulk of the energy of the photon candidate in the calorimeter is thus due to the brem'd photon. The candidates that enter our sample are the ones where the bending is so large that the track is no longer associated with the photon candidate by the standard photon reconstruction algorithms.

To understand this process in more detail we consider our $\gamma + \not{E}_T$ presample created from the set of W$\to e\nu$ MC events that have a photon in them that pass all the selection requirements in Table 5.3. Looking at the generator level information we can identify the location of the largest transfer of energy to a single photon [83]. We find that 93% of the time the electron gives more than 50% of its energy to a single photon in a single interaction. The remaining 7% are most likely tracking failures. Thus, we focus on hard-brem interactions as the primary cause of the $e \to \gamma_{fake}$ candidates. A simple requirement of fraction of the energy lost to be greater than 50% allows us to map out the locations of the hard bremsstrahlung interaction inside the detector. This is shown in the top of Figure 5.6. Note that this figure shows clearly the material inside the detector, an "x-ray" of sorts; showing the bulk of the bremsstrahlung interactions occur where the SVX detector and its support structure reside. The fact that so many of the electrons undergo bremsstrahlung early allows
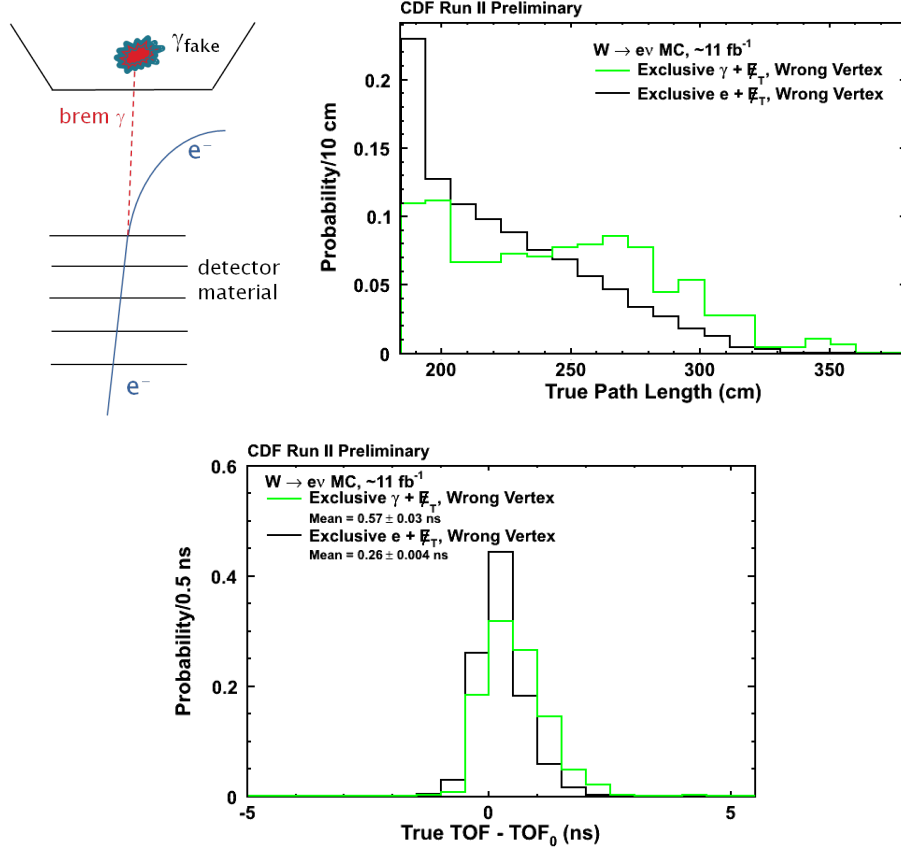
Fig. 5.5. (Top-LHS)A schematic representation of an electron interacting with the detector material and having a hard bremsstrahlung interaction. After the interaction the electron curves off because of its resulting lower energy and thus its trajectory becomes highly curved in the magnetic field and is no longer associated with the photon using the standard photon identification algorithms. It is important to note that both before and after the bremsstrahlung the trajectory can be reconstructed as a single low P_T track. (Top-RHS) The true path length for electrons mis-identified as photons, selected with Table 5.3 for the exclusive $\gamma + \not{E}_T$ presample, showing that these events tend to have larger path lengths than correctly identified electrons. (Bottom) The ΔTOF between the true vertex and $z = 0$ cm for the same sample of events demonstrating that events from electrons mis-identified as photons will have a larger bias.

us to understand why the conventional rejection methods fail, namely there is not much detector information available to reconstruct the track associated with the electron.
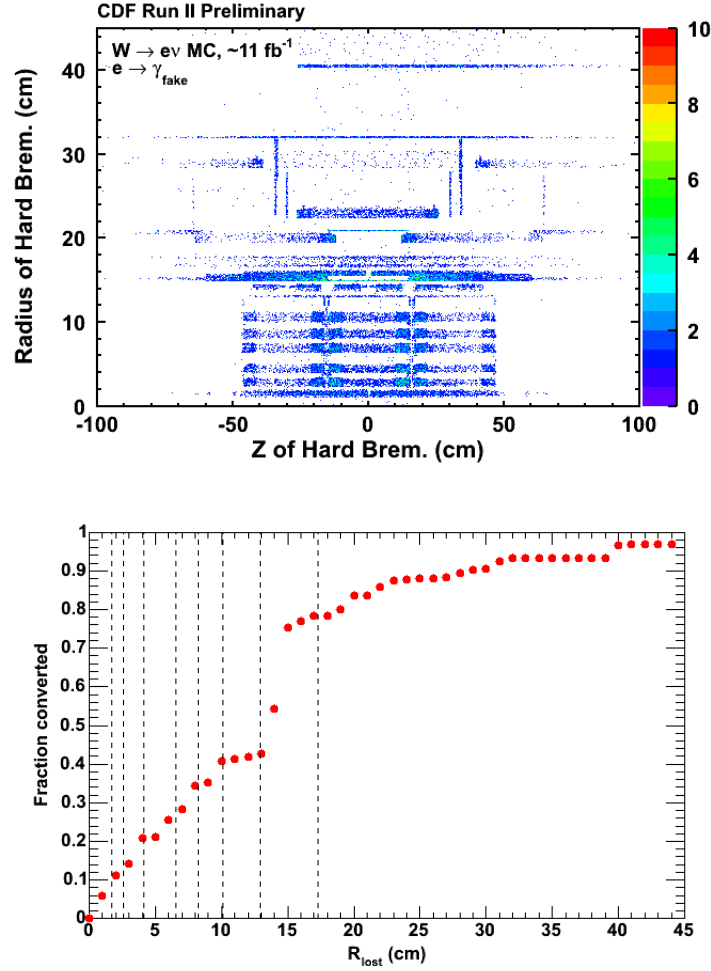

Fig. 5.6. Two different plots showing where, inside the detector, an electron interacts with the detector and loses more than half of its energy to a photon that is ultimately identified as passing all the photon ID requirements in Table 2.8. In the top plot we see a 2-D histogram showing the location in the radius vs. the z position along the beamline. The bottom shows an integral plot of the fraction of events which converted within the detector. Both indicate that the majority of events are seen to brem inside the silicon detector and the port cards (denoted with the dashed lines).

We can now also see why these photon candidates are not rejected by the N3D track requirement of the standard photon ID cuts given in Table 2.8. The post-brem electron loses its initial momentum and is left with low P_T, so the track is significantly curved away from the final location of the photon candidate in the calorimeter. Since the ϕ position of the low P_T track at the face of the calorimeter is far from the reconstructed ϕ position of the photon candidate, it is unlikely for the track to be "matched" to the photon candidate by the standard photon reconstruction algorithms [61]. We find that a low P_T brem'd track can end up at least three towers away from the EM cluster [83]. While the standard methods are effective for rejecting electrons in general, since the charged track is readily identified and rejected, we clearly can do much better.

Since the standard electron rejection methods leave a large number of fake events in our sample, and the remaining ones have a large time bias, we have developed a new method that takes advantage of the observation that the majority of the $e \rightarrow \gamma_{fake}$ candidates are due to electrons which interact with detector material and brem but still have a track that can be found. This method considers all reconstructed tracks that pass the requirements in Table 5.4 in the event and compares candidate tracks to the photon candidates in order to veto these events. We note that this track definition is a 'looser' definition then the tracks defined in Table 2.6 for 'good' timing tracks and high P_T isolated tracks in Table 5.4. The reason for this is we expect these tracks to only be present as high P_T tracks early in the detector and then undergo a hard brem changing their trajectory mid-flight, and thus be of lower quality. Therefore we use this looser definition in order to maximize the likelihood of finding this track.

Since the amount of variation in the expected measurement of the track position in ϕ and η are different, we do matching between the location of the photon in the calorimeter and the original track direction using a normalized variable. We define the matching variable ΔR_{Pull} to determine if the track is matched to the photon candidate as:

COTAxialSeg(5)	
Number of COT Axial Segments with hits	≥ 2
COTStereoSeg(5)	
Number of COT Stereo Segments with hits	≥ 2
$|z|$	
Z Position of the track	≤ 150 cm

Table 5.4
Track identification variables for use in $e \to \gamma_{fake}$ veto. Note, these variables are defined in Appendix B.1.

$$\Delta R_{Pull} = \sqrt{\Delta\phi_{Pull}^2 + \Delta\eta_{Pull}^2}. \tag{5.4}$$

where $\Delta\phi_{Pull}^2$ and $\Delta\eta_{Pull}^2$ are defined in order to account for the detector response as:

$$\Delta\phi_{Pull} = \frac{\Delta\phi}{\sigma_\phi} \tag{5.5}$$

where σ_ϕ is measured to be $\sigma_\phi = 8.1 \cdot 10^{-2}$ and

$$\Delta\eta_{Pull} = \frac{\Delta\eta}{\sigma_\eta} \tag{5.6}$$

and $\sigma_\eta = 6.3 \cdot 10^{-3}$ [83]. The top of Figure 5.7 shows the distribution of the closest track to the photon candidate in $\Delta\phi_{Pull}$ and $\Delta\eta_{Pull}$ as being symmetric and small for our control sample of W$\to e\nu \to \gamma_{fake}+ \not{E}_T$ events. This allows us to draw a circle in ΔR_{Pull} in order to veto events that are likely to have come from $e \to \gamma_{Fake}$ processes. The bottom of Figure 5.7 shows what the ΔR_{Pull} variable looks like for a presample of W$\to e\nu \to \gamma_{fake}+\not{E}_T$ (shown in Black) and a presample of $\gamma+\not{E}_T$ events from $Z\gamma \to \nu\nu\gamma$ events (shown in red) where both samples pass the requirements for the $\gamma+\not{E}_T$ presample requirements in Table 5.3.
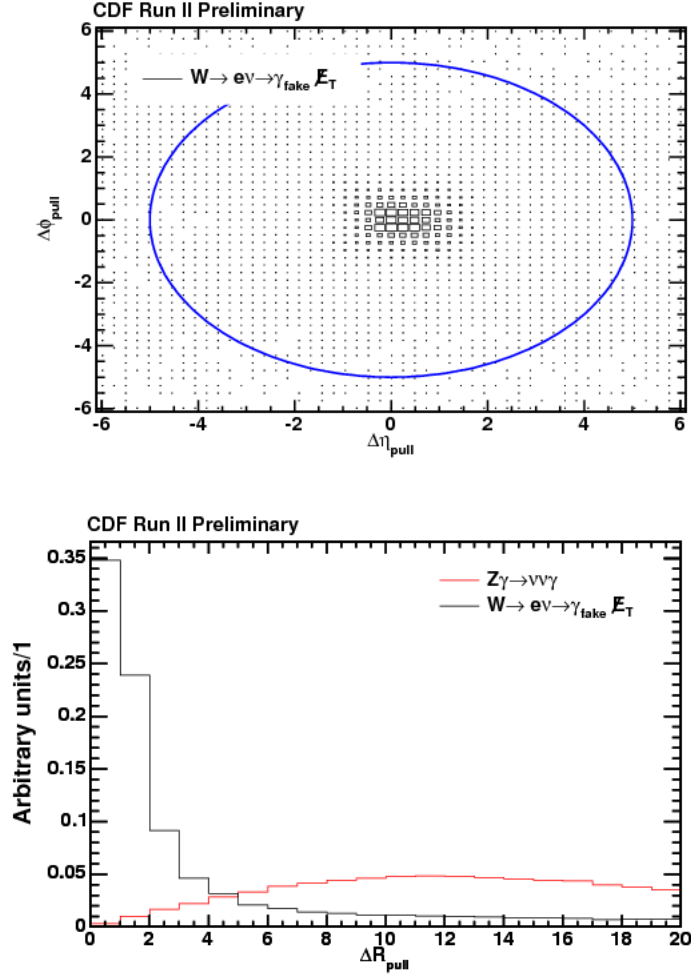
Fig. 5.7. To help reject electrons that fake photons, we have measured the angular separation between the photon and the closest track direction normalized to these measurement resolutions for our control sample of $\gamma+\not{E}_T$ events from W$\to e\nu$ MC with the added requirement that the photon come directly from an electron. The top plot shows the correlation between $\Delta\eta_{Pull}$ and $\Delta\phi_{Pull}$. The bottom plot shows a comparison of the ΔR_{pull} for our presample, along with a second $\gamma+\not{E}_T$ control sample from $Z\gamma \to \nu\nu\gamma$ MC showing the rejection power of this cut. Note, both samples are normalized unit area.

We have chosen to place a cut at $\Delta R_{Pull} > 5$ in order to veto $e \to \gamma_{fake}$ events. As shown in Figure 5.8, this selection is 93% efficient for real photons with a rejection power of 75% for $e \to \gamma_{fake}$. It is important to note, as shown in Figure 5.9, that this cut does not reduce the shifted mean of the wrong vertex distribution for $e \to \gamma_{fake}$ events. However it does reduce the overall rate at which they appear in our final sample. This reduces the overall importance of this background and makes us less sensitive to the wrong vertex mean shift.
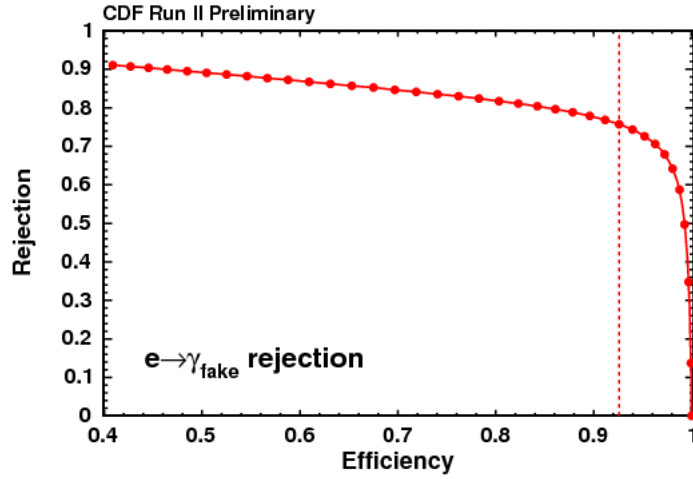
Fig. 5.8. This plot show the rejection of a electron rejection cut on ΔR_{pull} as a function of the efficiency. As the cut gets tighter the rejection gets worse but the efficiency goes up. A cut at $\Delta R_{Pull} = 5$ (red dashed line) results in approximately 93% efficiency of MC $Z\gamma \to \nu\nu\gamma \to \gamma + \not{E}_T$ and 75% rejection of $e \to \gamma_{fake}$.

5.5.3 Rejecting Events from Large $|Z_{Collision}|$ Sources

The final source of timing biases we address here comes from events where a collision occurs at $|z| > 60$ cm and produces a photon candidate that is then found in the calorimeter. Since vertices at $|z| > 60$ cm cannot be selected as the highest ΣP_T vertex in the exclusive $\gamma + \not{E}_T$ analysis, if a min-bias collision happens to occur
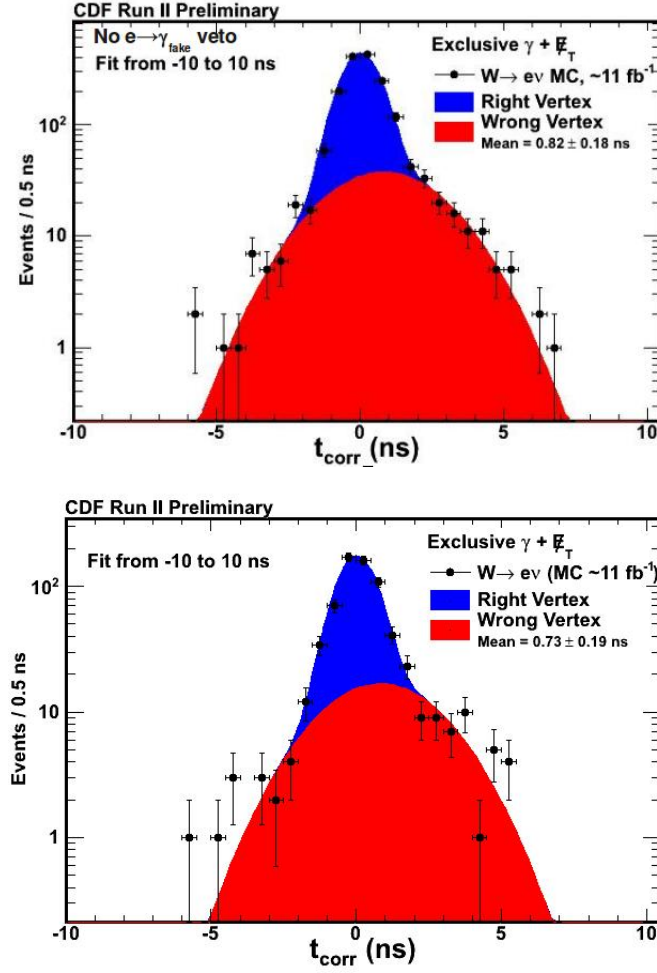
Fig. 5.9. The timing distribution for the $\gamma + \not{E}_T$ presample from W$\rightarrow e\nu$ MC before (top) and after (bottom) the ΔR_{Pull} cut. The application of this cut does not reduce the wrong vertex timing bias but does reduce the overall rate at which this background appears in our final sample.

at the center of the detector the only way this event is selected is if we mis-assign the vertex and thus incorrectly calculate the corrected time. Since these events always have $TOF_{RV} > TOF_{WV}$ they can have significantly biased $< t_{corr}^{WV} >$.

This situation is most easily seen in our $\gamma + \not{E}_T$ presample of γ + Jet events. In order for these QCD based events to enter the exclusive $\gamma + \not{E}_T$ they must have

a unique topology in order to produce fake $\not{E}_T$. In Figure 5.10 we plot the true z position of the collision for γ + Jet events simulated with MC and see that it extends far beyond $|z| = 60$ cm. Therefore we veto any event that has evidence there was a collision at >60 cm. Specifically, to be more efficient at rejecting these vertices, we use CDF's standard vertex algorithm, described in greater detail in reference [64] to search for vertices out to $|z| = 150$ cm and provide a handle on events that have evidence of activity at large collision z. If we find a standard vertex with three or more tracks at $|z| >60$ cm we veto this event as likely having a collision at large z position, as outlined in Table 2.11.
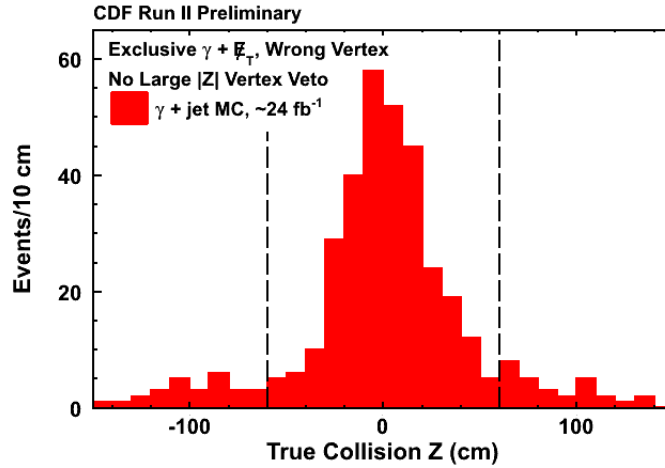

Fig. 5.10. The z distribution of the true collision position for a MC sample of γ+Jet events selected using Table 5.3 which defines the γ+$\not{E}_T$ presample. This distribution shows many events which originate at $|z| >60$ cm.

The effect of this veto can be seen in Figure 5.11 where we show the timing distribution of our γ+$\not{E}_T$ presample from γ + Jet MC. We then apply the large $|z|$ vertex veto to the sample and show that $< t^{WV}_{corr} >$ goes from 0.38 ns to 0.18 ns, greatly reducing the timing bias present in this sample.
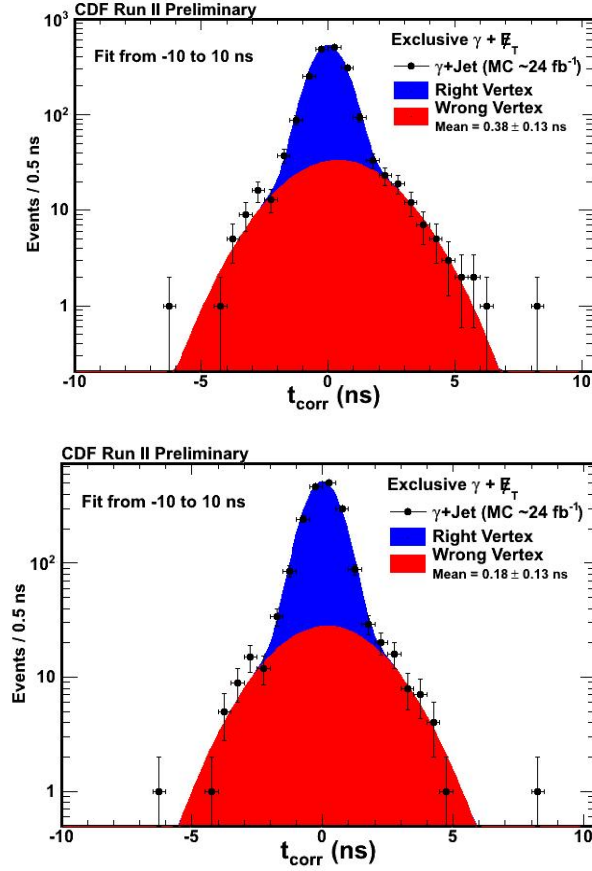
Fig. 5.11. The timing distribution for the $\gamma + \not{E}_T$ presample from γ+Jet MC events (Top) and the same sample after applying the large z veto (Bottom) showing the wrong vertex mean becomes much less biased.

Before we continue, we show that the large $|z|$ veto does not effect the timing distribution for a sample of events which originate inside the $|z| < 60$ cm area, for example as would be the case with a potential signal sample. To illustrate this, we consider again our $\gamma + \not{E}_T$ presample of $Z\gamma \rightarrow \nu\nu\gamma$ MC. In Figure 5.12 we show the t_{corr} distribution for the $Z\gamma$ events before and after the application the large $|z|$ veto showing very little effect to the timing distribution, as expected.
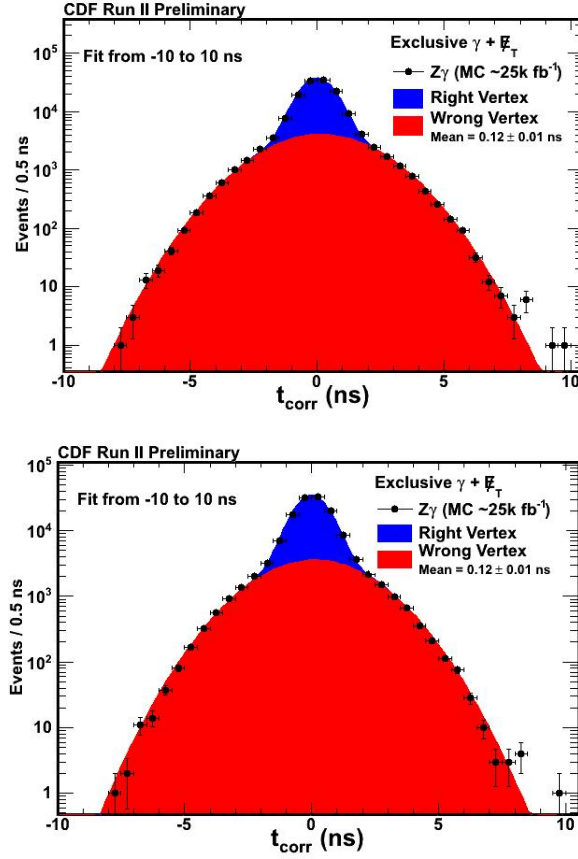
Fig. 5.12. The timing distribution for the $\gamma + \not{E}_T$ presample from from $Z\gamma \to \nu\nu\gamma$ MC events (Top) and the same sample after applying the large z veto (Bottom) showing very little effect in the timing distribution for events which originate from within $|z| < 60$ cm.

Finally, we estimate the efficiency of the large $|z|$ veto by applying it to our $\gamma + \not{E}_T$ presample, but where we select a subsample of cosmic ray events by considering events in the timing region 20 ns $< t_{corr} < 80$ ns. The timing distribution before and after the large $|z|$ veto shown in Figure 5.13. From this sample we estimate that the large $|z|$ veto is >95% efficient for real photons and collisions coming from $|z| < 60$ cm.
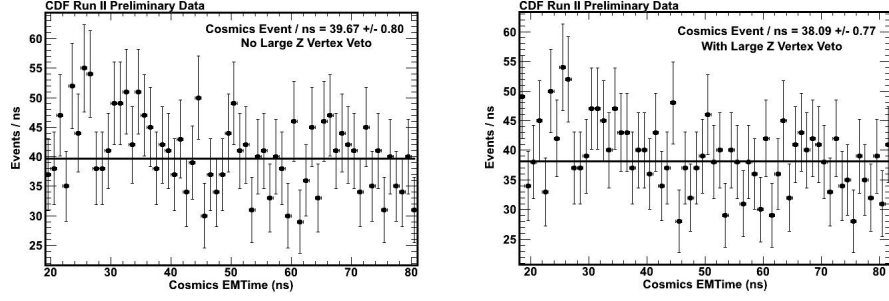
Fig. 5.13. The timing distribution for cosmic ray events selected using $\gamma + \not{E}_T$ data presample using the requirements in Table 5.3 but adding the restriction of the timing region from 20 ns $< t_{corr} <$ 80 ns. The left and right plots show the timing distribution before and after the large z veto the rate of cosmic rays is effectively not effected, as expected, by the large z veto

5.6   Timing Distributions for the Standard Model Backgrounds

Now that we have completed our discussion of the mechanisms for the production of SM events with large times, and methods for rejecting and/or minimizing the bias, we now consider the SM backgrounds as well as our electron samples after all the cuts. The final set of requirements are shown in Table 5.5.

With the final event selection established for the exclusive $\gamma + \not{E}_T$ final state we now select our control samples as passing the subset of events in our presample that also pass all the requirements in Tables 5.5 and 5.2. With these samples we can examine their timing distributions. Figure 5.14 shows the timing distribution for six SM MC background control samples as well as our two exclusive $e + \not{E}_T$ control samples in data using Tables 5.2 and 5.5. We see that each is well described by a double Gaussian fit where the right vertex mean and RMS are fixed to be 0.0 ns and 0.65 ns respectively and the normalization is allowed to float. Likewise the wrong vertex RMS is fixed to 2.0 ns while the mean and normalization of the distribution are allowed to float. The results are summarized in Table 5.6.
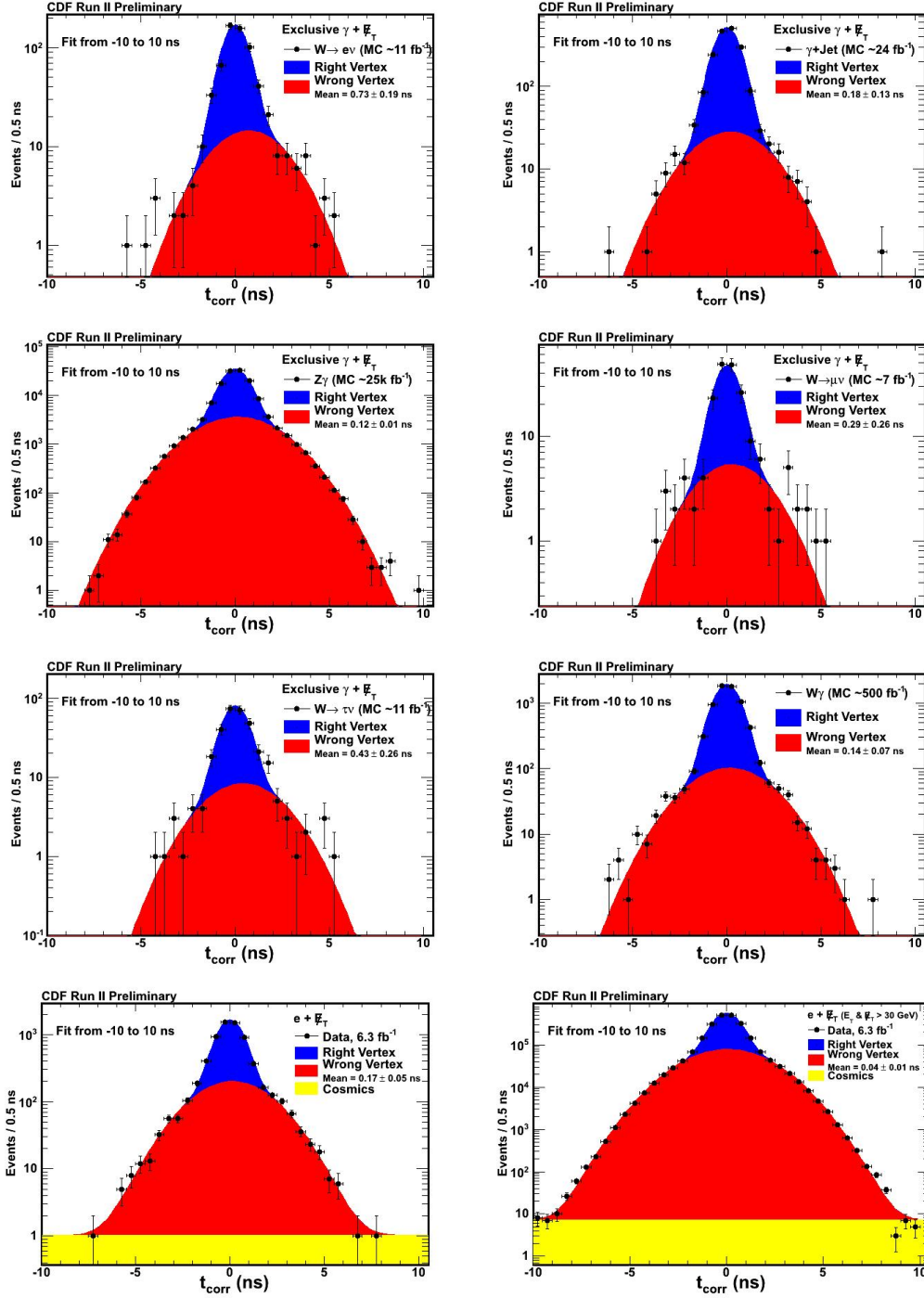
Fig. 5.14. The timing distributions for the W→ eν, γ+Jet, Zγ, W→ μν, W→ τν, and Wγ MC control samples and the e+$\not{E}_T$ control samples. The distributions are well fit by a double Gaussian distribution. In this fit the right vertex (blue) Gaussian is fixed with a mean =0.0 ns and a RMS =0.65 ns and the wrong vertex (red) Gaussian RMS is fixed =2.0 ns while the mean and normalization are allowed to vary.

Pass Trigger and Photon Good Run List
(See Table 2.2 and 2.3 and Section 2.4)
Pass Tight Photon requirements w/ $E_T^0 > 45$ GeV and $\not{E}_T^0 > 45$ GeV
(See Table 2.8 and Section 2.4.6)
Pass Beam Halo Rejection
(See Table 4.3)
Pass Cosmics Rejection
(See Table 4.2)
Pass Track Veto for Tracks with $P_T > 10$ GeV
(See Table 2.7)
Pass Jet Veto for Jets with $E_T^0 > 15$ GeV
(See Table 2.5)
Pass Large $|Z|$ Vertex Veto
(See Section 5.5.3)
Pass $e \rightarrow \gamma_{fake}$ Veto
(See Table 5.4 and Section 5.5.2)
Require a Good SpaceTime Vertex
(See Table 2.10)

Table 5.5

Exclusive $\gamma + \not{E}_T$ complete table of event selection requirements.

We can further examine one of the assumptions; namely the assumption that the wrong vertex distribution is described by a Gaussian with an RMS of 2.0 ± 0.1 ns. Figure 5.15 shows the results of testing this hypothesis for our six MC samples and two $e+ \not{E}_T$ data samples but where we allow the wrong vertex mean and RMS to vary for each sample during the fitting procedure and keep the right vertex mean and RMS fixed to 0.0 and 0.65 ns respectively. The results are summarized in Table 5.7. They clearly show that the assumption that the wrong vertex distribution is given by a Gaussian with RMS 2.0 ± 0.1 ns is an accurate description over a large range of wrong vertex means and across a variety of background samples.

With a complete set of data selection requirements well defined to reduce the bias and systematic production of large time events, and confidence that our SM collisions are well described by a double Gaussian with well understood RMS's but a variable

Sample	Wrong Vertex Mean (ns)
W→ eν MC	0.73 ± 0.19 ns
γ+Jet MC	0.18 ± 0.13 ns
Zγ MC	0.12 ± 0.01 ns
W→ $\mu\nu$ MC	0.29 ± 0.26 ns
W→ $\tau\nu$ MC	0.43 ± 0.26 ns
Wγ MC	0.14 ± 0.07 ns
$e+\not{E}_T$ Data	0.17 ± 0.05 ns
$e+\not{E}_T$ Data (E_T & $\not{E}_T$ > 30 GeV)	0.04 ± 0.01 ns

Table 5.6

Summary of the measured mean of the wrong vertex timing distribution, $< t^{WV}_{corr} >$, from the double Gaussian fit for our six MC and two $e+ \not{E}_T$ data control samples. Here the right vertex Gaussian is fixed with a mean =0.0 ns and a RMS =0.65 ns and the wrong vertex Gaussian RMS is fixed =2.0 ns while the mean and normalization are allowed to vary. These results are taken from Figure 5.14.

Sample	Wrong Vertex Mean (ns)	Wrong Vertex RMS (ns)
W→ eν MC	0.69 ± 0.22 ns	2.18 ± 0.17 ns
γ+Jet MC	0.18 ± 0.13 ns	2.04 ± 0.16 ns
Zγ MC	0.08 ± 0.05 ns	1.97 ± 0.05 ns
W→ $\mu\nu$ MC	0.30 ± 0.23 ns	2.06 ± 0.18 ns
W→ $\tau\nu$ MC	0.48 ± 0.22 ns	1.97 ± 0.22 ns
Wγ MC	0.14 ± 0.09 ns	2.14 ± 0.08 ns
$e+\not{E}_T$ Data	0.16 ± 0.07 ns	2.05 ± 0.07 ns
$e+\not{E}_T$ Data (E_T & $\not{E}_T$ > 30 GeV)	0.04 ± 0.05 ns	1.98 ± 0.05 ns

Table 5.7

Summary of the measured mean and RMS of the wrong vertex timing distributions, $< t^{WV}_{corr} >$ for our six MC $\gamma+\not{E}_T$ and two $e+\not{E}_T$ control samples where we allow the mean and RMS of the wrong vertex Gaussian to float in the fit. These results are plotted in Figure 5.15

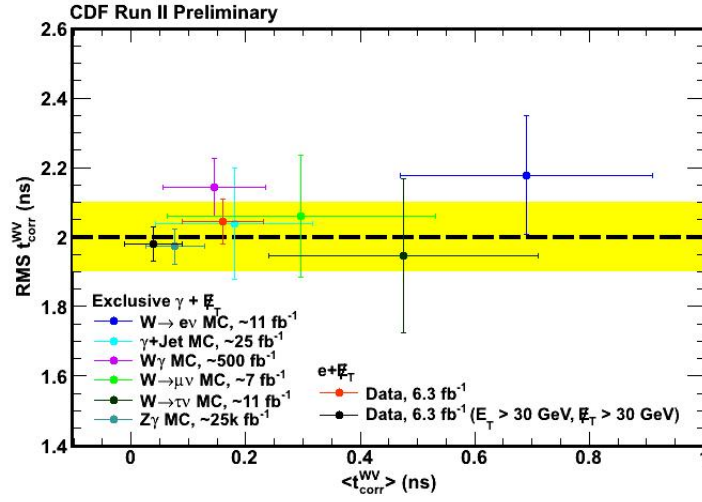**Fig. 5.15.** The measured mean RMS of the wrong vertex versus the mean of the wrong vertex timing distribution, $< t_{corr}^{WV} >$ where we have allowed both the mean and the RMS to vary in the fit for our six MC $\gamma + \not{E}_T$ and two $e + \not{E}_T$ data control samples. The results show that the description that the wrong vertex distribution is well modeled by a Gaussian with an RMS of 2.0 ± 0.1 ns.

WV mean, we next turn to the subject of being able to predict the wrong vertex mean which is needed to estimate the number of SM events in the signal region. In the next chapter, we will use events similar to those in the presamples to create a data-driven estimate of the number of background events in the signal region.